ASSESSING THE HABITABILITY OF EXOPLANETS: AN INTEGRATED FRAMEWORK INCORPORATING SIMILARITY EVALUATION AND KINETIC THEORY OF ATMOSPHERIC GASES

BACKGROUND

With the rapidly growing list of confirmed exoplanets (over 5,000, as shown in **Figure 1**) [1, 2], traditional methods for determining the habitability of exoplanets are becoming increasingly limited. Computational algorithms, commonly applied in machine learning classification and regression models and in novel data manipulation methods, are promising techniques for identifying potentially habitable exoplanets. Once identified, spectral analysis of atmospheric compositions of gases known to be biosignatures (such as oxygen, ozone, water vapor, carbon dioxide, methane, and nitrous oxide [3, 4]) can reveal insights into the potential for life on exoplanets and can be used to further determine the likelihood of habitability. However, remote detection and limited spatial and spectral information, as well as a lack of detailed models to interpret exoplanet atmospheric data, pose difficulties in studying these distant worlds. However, it is important to note that, while we cannot be certain of the existence of life, we can express the likelihood of habitability with a probability that considers established trends and observations. Additionally, there have been significant advancements in implementing novel technologies and analytical methods; thus, with currently available data, it is critical to prioritize potential candidates for future investigation. In this study, two frameworks were proposed and evaluated for determining the most Earth-like exoplanets.

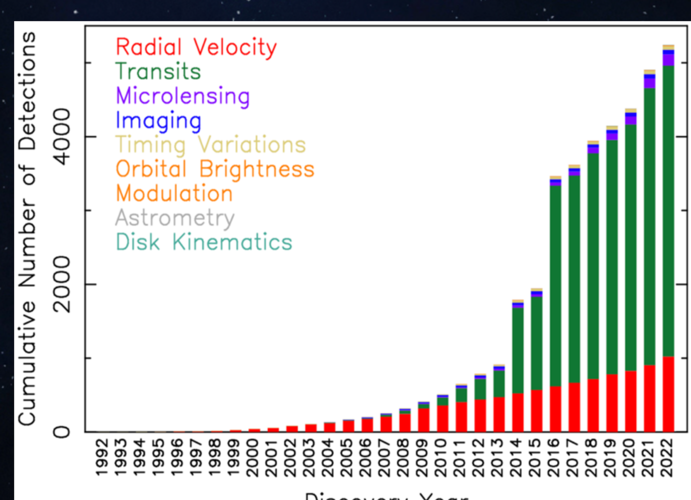

Figure 1. Graph shows the accumulated number of confirmed exoplanets for each year, with colors indicating the detection method for the exoplanets. [Image taken from ref. 2]

QUESTIONS/HYPOTHESIS

- What factors influence the possibility of life on exoplanets and how can they be taken into account when determining the likelihood of habitability?
- How can the evaluation of similarity through similarity metrics be applied to identify potentially habitable exoplanets and rank the most Earth-like ones?
- How can a thermal escape model be used to examine potential atmospheric biosignatures within each exoplanet's atmosphere?
- What are the most viable options for further investigation based on the quality and availability of data?

✓ Spectral analysis of exoplanet atmospheres can provide insights into the potential for life on exoplanets.
✓ Priority should be given to potential exoplanet candidates that have a high probability of habitability for future investigation.
✓ Advancements in technology and analytical methods could accomplish the following:
 - address the limitations in remote detection and the limited spatial and spectral information in exoplanet research.
 - effectively identify potentially habitable exoplanets.
✓ An integrated approach should be applied for future determination of exoplanet habitability.

METHODOLOGY

Frameworks for Assessing Habitability of Exoplanets

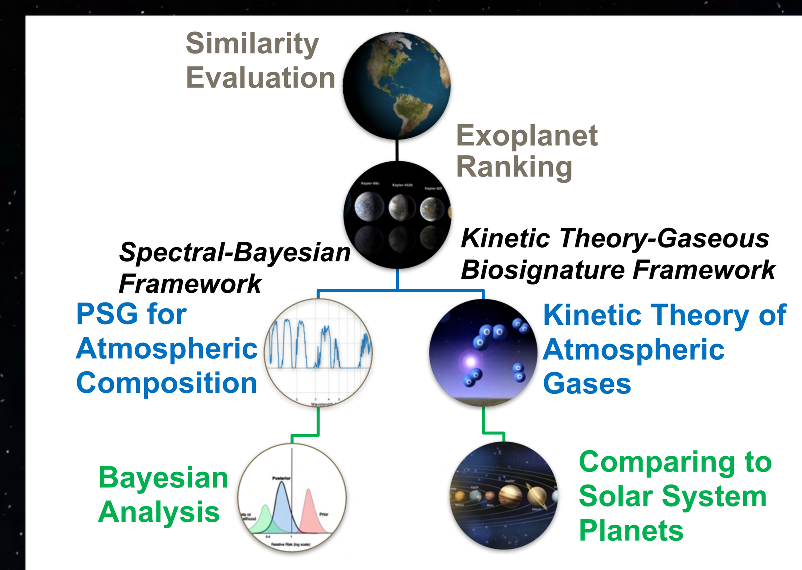

Figure 2. Overview of Spectral-Bayesian framework and Kinetic Theory-Gaseous Biosignature framework.

Similarity Evaluation Metrics

- Similarity metrics are comprised of a variety of statistical methods that can be used to quantify each exoplanets' habitability and how similar each is to Earth.
- All algorithms were coded in Python in the Google Colaboratory coding environment.

Similarity Evaluation Algorithms
Mean Squared Error (mse)
Euclidean Distance (euc)
Manhattan Distance (man)
Squared Euclidean Distance (sqe)
Canberra Distance (can)
Chebyshev Distance (chy)
Minkowski Distance (min)
Pearson Correlation Coefficient (pea)
Standardized Euclidean Distance (ste)
Bray-Curtis Dissimilarity (brc)
Earth Similarity Index (esi)

Datasets

- The first dataset consists of publicly accessible composition data for all confirmed exoplanets from the NASA Exoplanet Archive, Planetary Habitability Laboratory, and Planetary Spectrum Generator (PSG). [1, 8]
- The second dataset was specifically generated based on the computational results of exoplanets with higher scores, and includes data on escape velocity, thermal velocity, and biosignatures. Data for these same parameters for solar system planets were added in the second dataset as well.

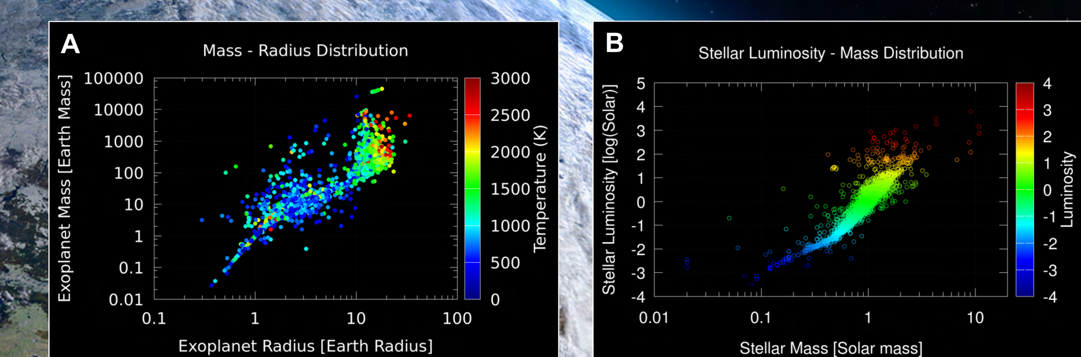

Figure 3. A) Mass and radius distribution for confirmed exoplanets. The colors show the temperatures for the corresponding exoplanets. B) Stellar luminosity and mass distribution for the confirmed planetary systems. The colors show the luminosities for the corresponding stars.

Parameters and Gaseous Biosignatures

	Exoplanets	Star
Parameters (Earth-like atmospheres) [1, 6, 8]	Star Type Number of Stars Number of Planets Orbital Period [Days] Planet Radius [Earth Radius] Planet Mass [Earth Mass] Planet Density [g/cm^3] Eccentricity Insolation Flux [Earth Flux] Equilibrium Temperature [K]	Effective Temperature [K] Radius [Solar Radius] Mass [Solar Mass] Luminosity [log(Solar)]
Atmospheric Gases (Gaseous biosignatures) [3, 4]	H_2, He, CH_4, NH_3, H_2O, HCN, CO, N_2, C_2H_6, O_2, H_2S, CO_2, N_2O, O_3, CH_3Cl, SO_2	

Thermal Escape Model of Gases

$$V_{esc} = \sqrt{2GM/R} \quad \textit{(eq. 1)}$$

$$V_{thermal} = \sqrt{3k_bT/m} \quad \textit{(eq. 2)}$$

$$T_p^4 = \frac{(1-a)}{4}\left(\frac{r}{d}\right)^2 T_{eff}^4 \quad \textit{(eq. 3)}$$

[5,9,10]

Parameters	
V_{esc}	Escape velocity (m/s)
$V_{thermal}$	Thermal velocity (m/s)
G	Universal gravitational constant
M	Mass of planet/exoplanet (kg)
R	Radius of planet/exoplanet (m)
T	Exosphere temperature (K)
k_b	Boltzmann's constant
m	Mass of one molecule of the gas (kg)
T_p	Calculated equilibrium temperature (K)
T_{eff}	Effective temperature of the star (K)
a	Bond albedo
r	Radius of the star (m)
d	Semi-major axis of the exoplanet (m)

RESULTS

1. Similarity Evaluation

- The top ten exoplanets were determined from each similarity method (**Table 1**).
- The data was then compared to Earth using Pearson's r correlation calculation, a measure of linear correlation. **(Figure 4)**
- In several graphs (Mean Squared Error, Manhattan Distance, Chebyshev Distance, Standardized Euclidean Distance, Pearson Correlation, and Composite), Earth's similarity to higher-ranked exoplanets is clear.
- Exoplanets ranked second through tenth are quite like one another.
- The graphs for the Canberra Distance, Standardized Euclidean Distance, and Earth Similarity Index similarity metrics differ more significantly from the other graphs because they produced relatively distinct rankings.
- Kepler-452 b, Kepler-69 c, and Kepler-1638 b are significant exoplanets in terms of similarity to Earth, and Kepler-452 b is the one that is most like Earth.

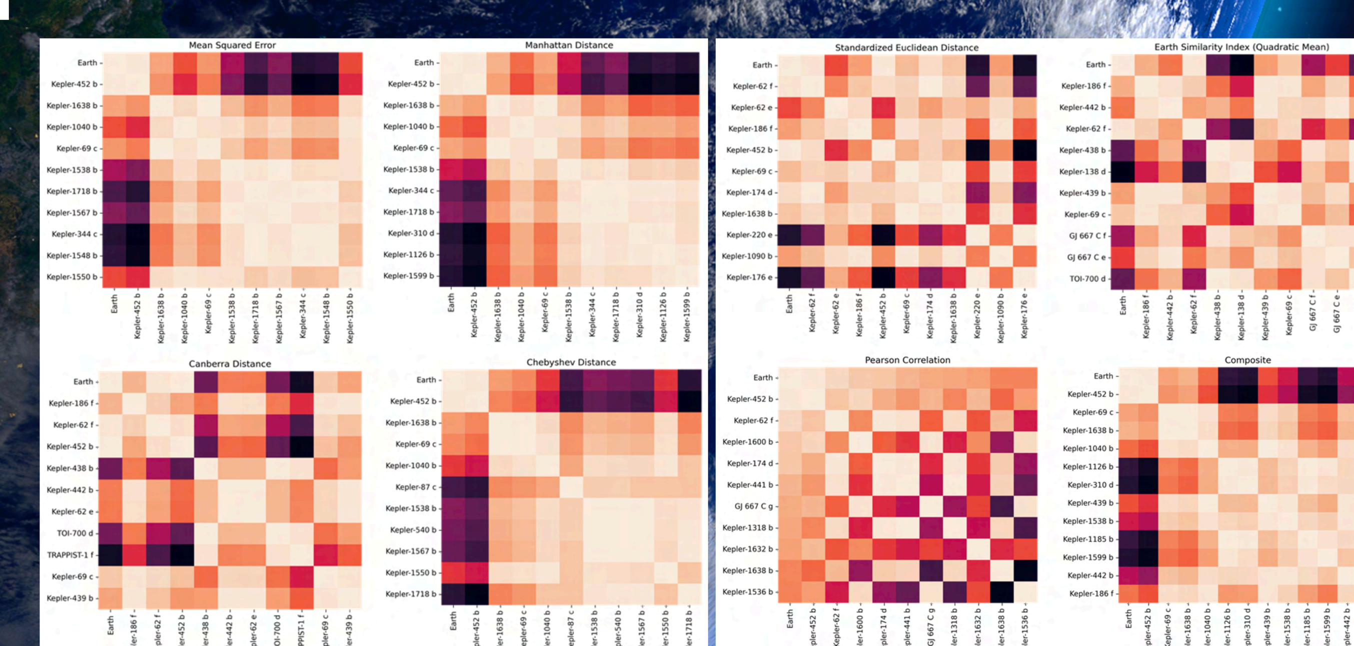

Figure 4. Representative results showing the top 10 most similar exoplanets to Earth, ranked by different similarity algorithms. Mean squared error, Euclidean distance, squared Euclidean distance, and Minkowski distance methods generated the same ranking. Manhattan distance and Bray-Curtis dissimilarity methods generated the same ranking.

Table 1. Exoplanets ranking based on similarity evaluations

Ranking	mse, euc, sqe, min	man, brc	can	chy	pea	ste	esi	Composite Ranking	Score
1	Kepler-452 b	Kepler-452 b	Kepler-186 f	Kepler-452 b	Kepler-452 b	Kepler-62 f	Kepler-186 f	**Kepler-452 b**	3
2	Kepler-1638 b	Kepler-1638 b	Kepler-62 f	Kepler-1638 b	Kepler-62 f	Kepler-62 e	Kepler-442 b	**Kepler-69 c**	5.45
3	Kepler-1040 b	Kepler-1040 b	Kepler-452 b	Kepler-69 c	Kepler-1600 b	Kepler-186 f	Kepler-62 f	**Kepler-1638 b**	8.55
4	Kepler-69 c	Kepler-69 c	Kepler-438 b	Kepler-1040 b	Kepler-174 d	Kepler-452 b	Kepler-438 b	**Kepler-1040 b**	12.55
5	Kepler-1538 b	Kepler-1538 b	Kepler-442 b	Kepler-87 c	Kepler-441 b	Kepler-69 c	Kepler-138 d	**Kepler-1126 b**	33.36
6	Kepler-1718 b	Kepler-344 c	Kepler-62 e	Kepler-1538 b	GJ 667 C g	Kepler-174 d	Kepler-439 b	**Kepler-310 d**	33.55
7	Kepler-1567 b	Kepler-1718 b	TOI-700 d	Kepler-540 b	Kepler-1318 b	Kepler-1638 b	Kepler-69 c	**Kepler-439 b**	34.55
8	Kepler-344 c	Kepler-310 d	TRAPPIST-1f	Kepler-1567 b	Kepler-1632 b	Kepler-220 e	GJ 667 C f	**Kepler-1538 b**	36.09
9	Kepler-1548 b	Kepler-1126 b	Kepler-69 c	Kepler-1550 b	Kepler-1638 b	Kepler-1090 b	GJ 667 C e	**Kepler-1185 b**	37.64
10	Kepler-1550 b	Kepler-1599 b	Kepler-439 b	Kepler-1718 b	Kepler-1536 b	Kepler-176 e	TOI-700 d	**Kepler-1599 b**	38.18

2. PSG Models of EARTH and VENUS

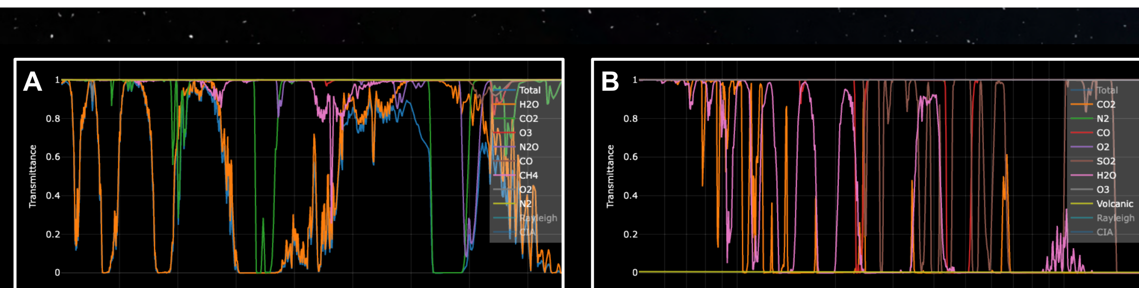

Figure 5. Representative transmission spectrum showing gaseous atmosphere of A) Earth ($N_2/O_2/H_2O$ planet) and B) Venus (CO_2 planet). The spectra were generated using the Planetary Spectrum Generator (PSG).

- The Planetary Spectrum Generator (PSG) is the platform used in this work to generate spectra and atmospheric compositions of exoplanets.
- PSG models can provide an accurate analysis for solar system planets.
- However, the PSG currently utilizes a terrestrial structure and composition for Earth-like exoplanets and a Venusian structure and composition for super-Earth like exoplanets, resulting in inaccurate spectra generation and atmospheric composition analysis for exoplanets. [6, 7]
- The Spectral-Bayesian framework is heavily dependent on the quality and availability of atmospheric compositions of exoplanets, which may not be readily available at this current stage.

RESULTS (Cont.)

3. Thermal Escape of Biosignatures

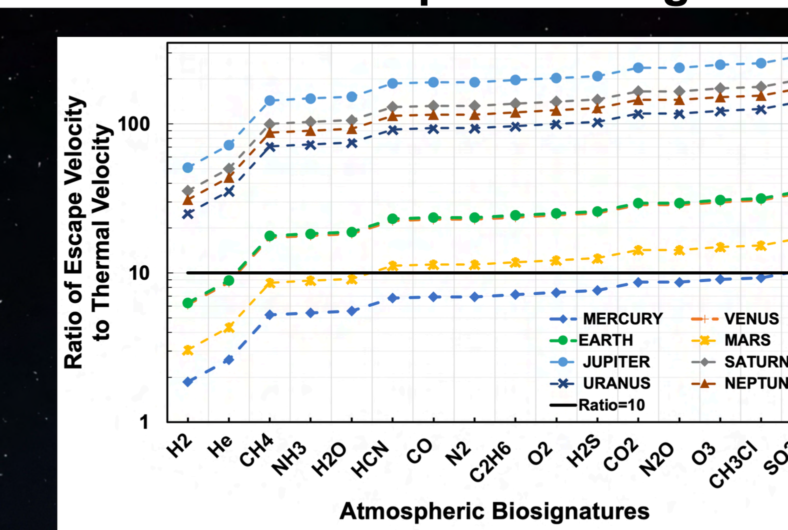

Figure 6. Ratio of escape velocity to thermal velocity of atmospheric species for the solar system planets. The solid horizontal line represents a ratio of 10 to 1.

Escaping gases for potential and known atmospheric biosignature gases from the solar system planets, based on the thermal escape model

Solar System Planets	Escaping Gases
Mercury	all major gases
Venus	H_2, He
Earth	H_2, He
Mars	H_2, He, CH_4, NH_3, H_2O
Giant planets (Jupiter, Saturn, Uranus, and Neptune)	no gases lost

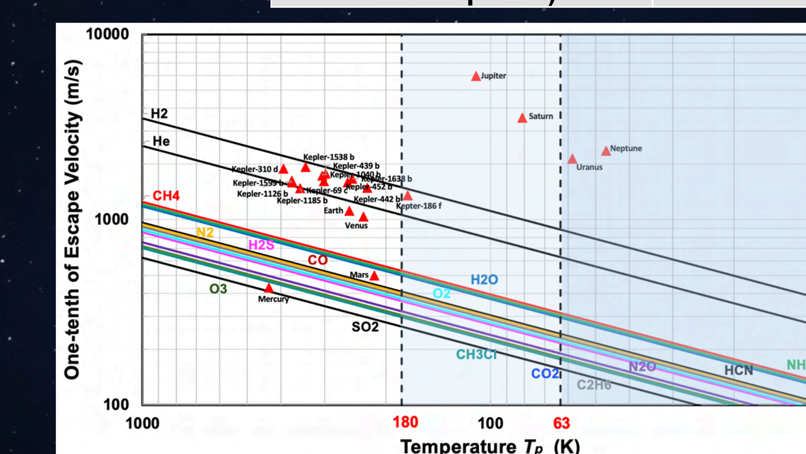

Figure 7. One-tenth of the escape velocity versus the equilibrium temperature T_p of the selected exoplanets and solar system planets. The lines represent the thermal velocities of various atmospheric biosignature gases. The vertical dashed lines (180K and 63K) indicate the freezing temperatures of H_2O and N_2, respectively, below which (the shaded right side of each line) the gases exist in the solid state.

CONCLUSION

Scientists have made significant progress in understanding the atmospheres and compositions of the planets in our solar system. However, exoplanets present a greater challenge, as the technology available today and the vast distances between us and these exoplanets make it difficult to gain detailed information about them. This makes it challenging to determine their habitability. For the Spectral-Bayesian approach to analysis, the framework's accuracy is heavily dependent on the quality and availability of atmospheric data. While PSG models can currently provide an accurate analysis for solar system planets, they are not suitable for exoplanets. Additionally, data from missions like the James Webb Space Telescope (JWST) may not be readily available, especially to the public. Thus, the second framework, which combines similarity evaluation and the thermal escape of atmospheric biosignatures, was found to be more feasible and could be used to study specific atmospheric gases and prioritize potential candidates for future investigation of the most Earth-like exoplanets. Thermal escape is an important process to consider when studying exoplanet habitability, as it can affect the long-term stability of an atmosphere and potentially lead to a loss of water and other volatiles that are important for the emergence and maintenance of life. This approach led to the selection of the top ten potentially habitable exoplanets and further analysis using the thermal escape model to identify possible retained atmospheric biosignatures. The results suggest these exoplanets could have Earth-like atmospheres and be capable of sustaining liquid water. This integrated method could significantly improve our understanding of exoplanet habitability, and future use of data from the JWST is a planned aspect of this project. While this integrated approach will greatly enhance our understanding of exoplanet habitability, it is important to note that only atmospheric biosignatures were considered in this work. Surface biosignatures also play a crucial role and should be considered in future comprehensive examinations of habitability to address the current limitations in determining the habitable zone of a star and the presence of an Earth-like atmosphere on an exoplanet.

REFERENCES

1. NASA Exoplanet Archive: Planetary Systems Composite Data, https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=PSCompPars.
2. NASA Exoplanet Archive: Cumulative Detections of Exoplanets https://exoplanetarchive.ipac.caltech.edu/exoplanetplots/exo_dischist_cumulative.png.
3. Catling, D.C., et al., Exoplanet Biosignatures: A Framework for Their Assessment. Astrobiology, 2018. 18(6): p. 709-738.
4. Schwieterman, E.W., et al., Exoplanet Biosignatures: A Review of Remotely Detectable Signs of Life. Astrobiology, 2018. 18(6): p. 663-708.
5. Konatham, S., J. Martin-Torres, and M.P. Zorzano, Atmospheric composition of exoplanets based on the thermal escape of gases and implications for habitability. Proc Math Phys Eng Sci, 2020. 476(2241): p. 20200148.
6. NASA Planetary Spectrum Generator, https://psg.gsfc.nasa.gov/.
7. Villanueva, G.L., et al., Planetary Spectrum Generator: An accurate online radiative transfer suite for atmospheres, comets, small bodies and exoplanets. Journal of Quantitative Spectroscopy and Radiative Transfer, 2018. 217: p. 86-104.
8. The Planetary Habitability Laboratory. https://phl.upr.edu/data.
9. Tian, F., Atmospheric Escape from Solar System Terrestrial Planets and Exoplanets. Annual Review of Earth and Planetary Sciences, 2015. 43(1): p. 459-476.
10. Madhusudhan, N., Exoplanetary Atmospheres: Key Insights, Challenges, and Prospects. Annual Review of Astronomy and Astrophysics, 2019. 57(1): p. 617-663.

Unless otherwise noted, all graphs and tables were created by the student.